

## Lower large deviations for geometric functionals

Christian Hirsch<sup>1</sup>, Benedikt Jahnel<sup>2</sup>, András Tóbiás<sup>3</sup>

submitted: October 15, 2019

<sup>1</sup> Universität Mannheim  
Mathematisches Institut  
68131 Mannheim  
Germany  
E-Mail: hirsch@uni-mannheim.de

<sup>2</sup> Weierstraß-Institut  
Mohrenstr. 39  
10117 Berlin  
Germany  
E-Mail: benedikt.jahnel@wias-berlin.de

<sup>3</sup> Technische Universität Berlin  
Institut für Mathematik  
Straße des 17. Juni 136  
10623 Berlin  
Germany  
E-Mail: tobias@math.tu-berlin.de

No. 2632  
Berlin 2019



---

2010 *Mathematics Subject Classification.* 60K35, 60F10, 82C22.

*Key words and phrases.* Large deviations, lower tails, stabilizing functionals, random geometric graph,  $k$ -nearest neighbor graph, relative neighborhood graph, Voronoi tessellation, clique count .

This work was co-funded by the German Research Foundation under Germany's Excellence Strategy MATH+: The Berlin Mathematics Research Center, EXC-2046/1 project ID: 390685689.

Edited by  
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)  
Leibniz-Institut im Forschungsverbund Berlin e. V.  
Mohrenstraße 39  
10117 Berlin  
Germany

Fax: +49 30 20372-303  
E-Mail: [preprint@wias-berlin.de](mailto:preprint@wias-berlin.de)  
World Wide Web: <http://www.wias-berlin.de/>

# Lower large deviations for geometric functionals

Christian Hirsch, Benedikt Jahnel, András Tóbiás

## Abstract

This work develops a methodology for analyzing large-deviation lower tails associated with geometric functionals computed on a homogeneous Poisson point process. The technique applies to characteristics expressed in terms of stabilizing score functions exhibiting suitable monotonicity properties. We apply our results to clique counts in the random geometric graph, intrinsic volumes of Poisson–Voronoi cells, as well as power-weighted edge lengths in the random geometric,  $k$ -nearest neighbor and relative neighborhood graph.

## Contents

<b>1</b>	<b>Introduction and main results</b>	<b>1</b>
<b>2</b>	<b>Examples</b>	<b>4</b>
2.1	Clique counts and power-weighted edge lengths in random geometric graphs . . . . .	4
2.2	Intrinsic volumes of Voronoi cells . . . . .	5
2.3	Power-weighted edge counts in $k$ -nearest neighbor graphs and relative neighborhood graphs . . . . .	6
<b>3</b>	<b>Proofs</b>	<b>8</b>
3.1	Proof of Theorem 1.1 . . . . .	8
3.2	Proof of Theorem 1.2 . . . . .	8
3.3	Proof of Theorem 1.3 . . . . .	10

## 1 Introduction and main results

Considering the field of random graphs, there is a subtle difference in the understanding between upper and lower tails in a large-deviation regime. For instance, when considering the triangle count in the Erdős–Rényi graph, the probability of observing atypically few triangles is described accurately via very general Poisson-approximation results [Jan90, JW16]. On the other hand, the probability of having too many triangles requires a substantially more specialized and refined analysis [Cha12].

This begs the question whether a similar dichotomy also arises in the large-deviation analysis of functionals that are of geometric rather than combinatorial nature. For instance, Figure 1.1 shows a typical realization of the random geometric graph in comparison to a realization with an atypically small number of edges. In geometric probability, elaborate results are available for large and moderate

deviations of geometric functionals exhibiting a similar behavior in the upper and the lower tails [SY01, SY05, ERS15]. However, they prominently do not cover the edge count in the random geometric graph, whose upper tails have been understood only recently [CH14].

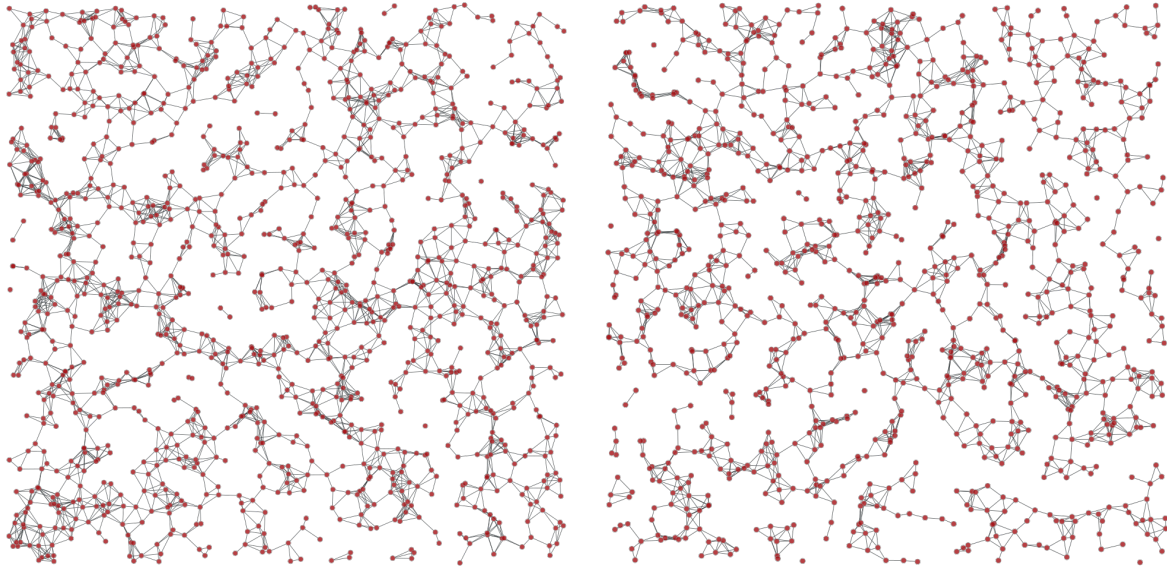


Figure 1.1: Typical realization of the random geometric graph (left) next to a realization having fewer than 75% of the expected number of edges (right).

In the present work, we provide three general results, Theorems 1.1, 1.2 and 1.3, tailored to studying large-deviation lower tails of geometric functionals. For the proofs, we resort to a method inspired by the idea of sprinkling [ACC<sup>+</sup>83]. We perform small changes in those parts of the domain where the underlying point process exhibits highly pathological configurations. After this procedure, we can compare the resulting functionals to approximations that are then amenable to the point-process based large-deviation theory from [GZ93] or [SY01, SY05]. Among the examples covered by our method are clique counts in the random geometric graph, inner volumes of Poisson–Voronoi cells and power-weighted edge lengths in the random geometric,  $k$ -nearest neighbor and relative neighborhood graph.

In the rest of this section, we set up the notation and state the main results. Then, Section 2 illustrates those results through the examples. Finally, Section 3 contains the proofs.

We study functionals on a homogeneous Poisson point process  $X = \{X_i\}_{i \geq 1} \subset \mathbb{R}^d$  with intensity 1, whose distribution on the space  $\mathbf{N}$  of locally-finite configurations will be denoted by  $\mathbb{P}$ . Following the framework of [SY01], these functionals are realized as averages of scores associated to the points of  $X$ . More precisely, a *score function*

$$\xi : \mathbb{R}^d \times \mathbf{N} \rightarrow [0, \infty)$$

is any bounded measurable function. To simplify notation, we shift the coordinate system to the considered point and write  $\xi(X - X_i) = \xi(X_i, X)$ . In this notation  $\varphi \mapsto \xi(\varphi)$  acts on configurations  $\varphi \in \mathbf{N}_o$ , the family of locally-finite point configurations with a distinguished node at the origin  $o \in \mathbb{R}^d$ .

We then consider lower tails of functionals of the form

$$H_n = H_n^\xi(X) = \frac{1}{n^d} \sum_{X_i \in X \cap Q_n} \xi(X - X_i), \quad (1.1)$$

i.e., averages of the score function over all points in the box  $Q_n = [-n/2, n/2]^d$  of side length  $n \geq 1$  centered at the origin.

In a first step, we derive upper bounds for the lower tail probabilities. To that end, we work with approximating score functions  $\xi^r$  that are  $r$ -dependent for some  $r > 0$ . That is,  $\xi^r(\varphi) = \xi^r(\varphi \cap B_r)$  for every  $\varphi \in \mathbf{N}_o$ , where  $B_r$  denotes the Euclidean ball of radius  $r$  centered at the origin.

To state the main results, we resort to the entropy-based formulation of the large-deviation rate function. We write

$$h(\mathbb{Q}) = \lim_{n \uparrow \infty} \frac{1}{n^d} \int d\mathbb{Q}_n \log \frac{d\mathbb{Q}_n}{d\mathbb{P}_n}$$

for the *specific relative entropy* of a stationary point process  $\mathbb{Q}$ , where  $\mathbb{Q}_n$  and  $\mathbb{P}_n$  denote the restrictions of  $\mathbb{Q}$  and  $\mathbb{P}$  to the box  $Q_n$ , respectively. If  $\mathbb{Q}_n$  is not absolutely continuous with respect to the restricted Poisson point process, we adhere to the convention that the above integral is infinite. Further,  $\mathbb{Q}^o[\xi]$  is the expectation of  $\xi$  with respect to the Palm version  $\mathbb{Q}^o$  of  $\mathbb{Q}$ , see [GZ93] for details. Here is our first main theorem.

**Theorem 1.1** (Upper bound). *Let  $a > 0$  and assume the score function  $\xi$  to be the pointwise increasing limit of a family  $\{\xi^r\}_{r \geq 1}$  of  $r$ -dependent score functions. Then,*

$$\limsup_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n \leq a) \leq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi] \leq a} h(\mathbb{Q}). \quad (1.2)$$

For the lower bound, we give two sets of conditions. The first deals with score functions  $\xi$  that are *increasing* in the sense that  $\xi(\varphi) \leq \xi(\psi)$  for every  $\varphi \subset \psi$ . This applies for instance to clique counts and power-weighted edge lengths in the random geometric graph.

**Theorem 1.2** (Lower bound for bounded-range scores). *Let  $a > 0$  and assume the score function  $\xi$  to be increasing and  $r$ -dependent for some  $r > 0$ . Moreover, assume that for every  $b > 0$  there exists  $M = M(b) > 0$  such that  $\xi(\varphi) \leq M$  whenever  $\#\varphi < b$ . Then,*

$$\liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n < a) \geq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi] < a} h(\mathbb{Q}). \quad (1.3)$$

However, many score functions are neither  $r$ -dependent nor increasing, or not even monotone. A prime example is the sum of power-weighted edge lengths in the  $k$ -nearest neighbor graph, see Section 2. Still, this example and many other score functions are stabilizing,  $R$ -bounded and weakly decreasing in the following sense.

First, a score function  $\xi$  is *stabilizing* if there exists a  $\mathbb{P}^o$ -almost surely finite measurable *stabilization radius*  $R: \mathbf{N}_o \rightarrow [0, \infty]$ , such that  $\{R(X) \leq r\}$  is measurable with respect to  $X \cap B_r$  for every  $r \geq 0$  and

$$\mathbb{P}^o(\xi(X) = \xi(X \cap B_{R(X)})) = 1.$$

In words,  $\xi(X)$  does not depend on the configuration outside the ball  $B_{R(X)}$ . We call  $R$  *decreasing* if  $R(\varphi \cup \{x\}) \leq R(\varphi)$  for all  $\varphi \in \mathbf{N}_o$  and  $x \in \mathbb{R}^d$ .

Second,  $\xi$  is  *$R$ -bounded* if for every  $\delta > 0$  and sufficiently large  $M = M(\delta) \geq 1$ ,

$$\mathbb{P}^o(\{R(X) \leq M\} \cap \{\xi(X) \geq \delta M^d\}) = 0.$$

Loosely speaking, the score function is negligible compared to the  $d$ th power of the stabilization radius.

Third,  $\xi$  is *weakly decreasing* if

$$\mathbb{P}(\#\{y \in X : \xi(X \cup \{o\} - y) > \xi(X - y)\} \leq k) = 1$$

holds for some  $k \geq 1$ . In words, for all but at most  $k$  points of a configuration, adding a new point to the configuration decreases the score function value of the point.

Finally, we need to ensure that sprinkling a sparse configuration of Poisson points yields control on the stabilization radii of the points in a box. More precisely, we assume that the stabilization radius is *regular* in the following sense. Let  $X^{+,M}$  denote a Poisson point process with intensity  $M^{-d}$  that is independent of  $X$ . Then, we assume that there exists  $K_0 > 0$  with the following property. For every  $\delta > 0$  there exist  $M_0 = M_0(\delta) \geq 1$  and  $n_0 = n_0(\delta) \geq 1$  such that for all  $M \geq M_0$  and  $n \geq n_0$ ,

$$\mathbb{P}(\{X^{+,M}(Q_n) \leq K_0(n/M)^d\} \cap E_n^{M,+} | X) \geq \exp(-\delta n^d)$$

holds almost surely. Here, for  $\varphi \in \mathbb{N}$  and any measurable subset  $A \subset \mathbb{R}^d$ , we write  $\varphi(A) = \#\{x \in \varphi : x \in A\}$  for the number of points of  $\varphi$  contained in  $A$ , and

$$E_n^{M,+} = \max_{X_i \in (X \cup X^{+,M}) \cap Q_n} R((X \cup X^{+,M}) - X_i) \leq M$$

denotes the event that after the sprinkling, the stabilization radii of all points in  $Q_n$  are at most  $M$ . Here is the corresponding main result.

**Theorem 1.3** (Lower bound for stabilizing scores). *Let  $a > 0$  and  $\xi$  be a weakly-decreasing  $R$ -bounded score function with a decreasing and regular radius of stabilization. Then, (1.3) remains true.*

## 2 Examples

In this section, we discuss how to apply the results announced in Section 1 to a variety of examples arising in geometric probability. More precisely, Sections 2.1, 2.2 and 2.3 are devoted to characteristics for the random geometric graph, the Voronoi tessellation,  $k$ -nearest neighbor graphs and relative neighborhood graphs, respectively.

### 2.1 Clique counts and power-weighted edge lengths in random geometric graphs

As a first simple application of our results, consider the set

$$C_k(\varphi) = C_{k,t}(\varphi) = \#\{x_1, \dots, x_k\} \subset \varphi : x_1 = o \text{ and } |x_i - x_j| < t \text{ for all } i \neq j\}$$

of  $k$ -cliques associated to the origin in the geometric graph on  $\varphi \in \mathbb{N}_o$  with connectivity radius  $t > 0$ . Then, for  $k \geq 2$  and  $\alpha \geq 0$ , the score functions

$$\xi_k(\varphi) = \frac{1}{k} \#C_k(\varphi) \quad \text{and} \quad \xi'_\alpha(\varphi) = \frac{1}{2} \sum_{x \in \varphi : |x| < t} |x|^\alpha$$

count the number of  $k$ -cliques containing the origin and the power-weighted edge lengths at the origin, respectively. Note that  $\xi_k$  and  $\xi'_\alpha$  are  $t$ -dependent and increasing. Additionally, if  $\#\varphi < b$ , then  $\xi_k(\varphi) \leq k^{-1}b^{k-1}$  and  $\xi'_\alpha(\varphi) \leq t^\alpha b$ . Hence Theorems 1.1 and 1.2 are applicable.

Further examples arise in the context of topological data analysis. More precisely, the number of  $k$ -cliques containing the origin is precisely the number of  $k$ -simplices of the Vietoris–Rips complex containing the origin. Similar arguments also apply to the Čech complex, the second central simplicial complex in topological data analysis. We refer the reader to [BCY18, Section 2.5] for precise definitions and further properties.

## 2.2 Intrinsic volumes of Voronoi cells

Recall the definition of the Voronoi cell at the origin of a locally-finite configuration  $\varphi \in \mathbf{N}_o$ , i.e.,

$$C_o(\varphi) = \{x \in \mathbb{R}^d : |x| \leq \inf_{y \in \varphi} |x - y|\}.$$

Recall that since  $C_o(\varphi)$  is a convex body, its *intrinsic volumes*  $v_0(C_o), v_1(C_o), \dots, v_d(C_o)$  can be computed. They are key characteristics of a convex set, e.g.,  $v_1, v_{d-1}$  and  $v_d$  are proportional to the mean width, the surface area and the volume, respectively. We refer the reader to [SW08, Section 14.2] for a precise definition and further properties. In particular, considering  $v_1$  in dimension  $d = 2$ , the associated characteristic  $n^d H_n$  becomes the total edge length of the Voronoi graph, so that we obtain a link to the setting studied in [SY05, Section 2.4.1]. Due to the intricate geometry, deriving a full large deviation principle even for a strictly concave function of the edge length was only achieved for a Poisson point process that is restricted to a lattice instead of living in the entire Euclidean space. This example illustrates that even in situations where understanding the large-deviation upper tails requires a delicate geometric analysis, the lower tails may be more accessible.

More precisely, consider the score functions

$$\xi_k(\varphi) = v_k(C_o(\varphi))$$

and note that  $\xi_k^r(\varphi) = v_k(C_o(\varphi) \cap B_r)$  is a  $4r$ -dependent, pointwise increasing approximation of  $\xi_k(\varphi)$ . Hence, the upper bound of Theorem 1.1 applies.

For the lower bound, the conditions of Theorem 1.3 can be satisfied using the following definitions. The radius of stabilization is described in [Pen07, Section 6.3]: Take any collection  $\{S_i\}_{i \in I}$  of cones with apex at the origin and angular radius  $\pi/12$  whose union covers  $\mathbb{R}^d$ , where  $I = I(d) \in \mathbb{N}$ . Let  $S_i^+$  denote the cone that has the same apex and symmetry hyperplane as  $S_i$  and has the larger angular radius  $\pi/6$ . Then, we define the stabilization radius

$$R(\varphi) = 2 \max_{i \in I} \min_{x \in \varphi \cap S_i^+} |x|, \quad (2.1)$$

as twice the radius at which the origin has a neighbor in every extended cone. In particular, both  $R$  and  $\xi_k$  are decreasing. Since  $C_o(\varphi) \subset B_{R(\varphi)}$ , we deduce that

$$\xi_k(\varphi) \leq v_k(B_{R(\varphi)}) = R(\varphi)^k v_k(B_1).$$

In particular,  $\xi_k$  is  $R$ -bounded for  $k < d$ . Finally, we define for a suitable constant  $L = L(d) \geq 1$  the event

$$A_n^M = \{X^{+,M}(Q_{M/L}(z)) = 1 \text{ for all } z \in (M/L)\mathbb{Z}^d \cap Q_{2n}\} \quad (2.2)$$

that  $X^{+,M}$  has precisely one point in each sub-box from an  $M/L$ -partition of the box  $Q_{2n}$ . It follows from the definition of  $R$  that the event  $E_n^{M,+}$  occurs whenever  $A_n^M$  occurs, provided that  $L$  is chosen sufficiently large. Moreover, setting  $K_0 = (2L)^d$ , we deduce that  $X^{+,M}(Q_n) \leq K_0(n/M)^d$  under

$A_n^M$ . Hence, it remains to establish the asserted lower bound on the probability  $\mathbb{P}(A_n^M)$ . Fixing  $\delta > 0$  and invoking the independence property of the Poisson point process yields that

$$\mathbb{P}(A_n^M) = \mathbb{P}(X^{+,M}(Q_{M/L}) = 1)^{(2nL/M)^d} = e^{-(2n/M)^d} L^{-(2nL/M)^d} \geq e^{-\delta n^d},$$

provided that  $M = M(\delta)$  is sufficiently large. Summarizing the above findings, we deduce that Theorem 1.3 can be applied to get the lower bound on the rate function.

## 2.3 Power-weighted edge counts in $k$ -nearest neighbor graphs and relative neighborhood graphs

Finally, we elucidate how to apply Theorem 1.3 to the power-weighted edge count of two central graphs in computational geometry, namely the  $k$ -nearest neighbor graph and the relative neighborhood graph. As we shall see, in contrast to the Voronoi example presented in Section 2.2, we encounter here score functions that are weakly decreasing but not decreasing. A full large deviation principle for the total edge length of the  $k$ -nearest neighbor graph is described in [SY05, Section 2.3], and we believe that the proof should extend to power-weighted edge lengths with a power strictly less than  $d$ . Nevertheless, we apply here our approach towards the large-deviation lower tails as it can be directly adapted to the bidirectional  $k$ -nearest neighbor graph, the relative neighborhood graph and possibly further graphs.

In the *undirected  $k$ -nearest neighbor graph*,  $\xi$  expresses the powers of distances between any point and the origin, such that at least one of them belongs to the set of  $k$  nearest neighbors of the other one. To be more precise,

$$\mathfrak{R}_k(\varphi) = \inf\{r > 0: \varphi(B_r) \geq k + 1\} \quad (2.3)$$

defines the  *$k$ -nearest neighbor radius* of  $o$  in  $\varphi \in \mathbf{N}_o$ . Then, for some  $\alpha \geq 0$ , the score function corresponding to the sum of power-weighted edge lengths of the  $k$ -nearest neighbor graph is defined via

$$\xi_{k,\alpha}(\varphi) = \frac{1}{2} \sum_{x \in \varphi: |x| \leq \mathfrak{R}_k(\varphi) \vee \mathfrak{R}_k(\varphi - x)} |x|^\alpha.$$

In particular, we recover the number of edges by setting  $\alpha = 0$ . As noted in [Pen07, Section 6.3], to construct a radius of stabilization we can proceed as in (2.1) except for replacing  $\min_{x \in \varphi \cap S_i^+} |x|$  by the distance of the  $k$ th closest point from the origin in  $\varphi \cap S_i^+$ . Hence,  $\xi_{k,\alpha}$  becomes stabilizing with a decreasing stabilization radius. In the same vein, a minor adaptation of the arguments in Section 2.2 yield the regularity and  $R$ -boundedness for  $\alpha < d$ .

In order to apply Theorem 1.3 for the lower bound, it remains to verify the following.

**Lemma 2.1.**  $\xi_{k,\alpha}$  is weakly decreasing.

*Proof.* Let us call  $\varphi \in \mathbf{N}$  *nonequidistant* if for all  $y, z, v, w \in \varphi$ ,  $|y - z| = |v - w| > 0$  implies  $\{y, z\} = \{v, w\}$ . First note that for any  $x \in \mathbb{R}^d$ , under  $\mathbb{P}$ , almost all configurations  $\varphi \cup \{x\}$  are nonequidistant. We claim that for any nonequidistant configuration  $\varphi \cup \{x\}$ , we have for all but at most  $k$  points  $y \in \varphi$  that

$$\xi_k(\varphi \cup \{x\} - y) \leq \xi_k(\varphi - y). \quad (2.4)$$



Indeed, for  $y \in \varphi$ , let us define the set of  $k$  nearest neighbors of  $y$  in  $\varphi$  as follows

$$k\text{NN}(\varphi, y) = (B_{\mathfrak{R}_k(\varphi-y)}(y) \cap \varphi) \setminus \{y\}.$$

Now, if  $y \in k\text{NN}(\varphi \cup \{x\}, x)$ , then possibly  $\xi_k(\varphi \cup \{x\} - y) > \xi_k(\varphi - y)$ . We claim that else (2.4) holds. Indeed, if  $y \notin k\text{NN}(\varphi \cup \{x\}, x)$ , then there are two possibilities. If  $x \in k\text{NN}(\varphi \cup \{x\}, y)$ , then  $x$  replaced precisely one neighbor  $z$  of  $y$  and is closer to  $y$  than  $z$ . More precisely, note that  $|x - y| \leq \mathfrak{R}_k(\varphi \cup \{x\} - y) \leq \mathfrak{R}_k(\varphi - y)$ . Hence, there exists  $z \in k\text{NN}(\varphi, y)$  such that  $|z - y| = \mathfrak{R}_k(\varphi - y)$  and  $z \notin k\text{NN}(\varphi \cup \{x\}, y)$ , the neighbor of  $y$  that is replaced by  $x$ . Additionally, for any  $w \in k\text{NN}(\varphi, y) \setminus \{z\}$  also  $w \in k\text{NN}(\varphi \cup \{x\}, y)$ . Further, also for any  $v \in \varphi$  such that  $y \in k\text{NN}(\varphi \cup \{x\}, v)$  we have  $y \in k\text{NN}(\varphi, v)$ . Hence,

$$\xi_k(\varphi \cup \{x\} - y) - \xi_k(\varphi - y) \leq |x - y|^\alpha - |z - y|^\alpha \leq 0,$$

which is (2.4). The other possibility is that  $x \notin k\text{NN}(\varphi \cup \{x\}, y)$ . Then the addition of  $x$  can only remove edges that were present due to the fact that some other point had  $y$  as a neighbor. In this case,  $\xi(\varphi \cup \{x\} - y) = \xi(\varphi - y)$  unless there exists  $z \in \varphi$  such that  $y \in k\text{NN}(\varphi, z)$  but  $y \notin k\text{NN}(\varphi \cup \{x\}, z)$ , which must be due to the property that  $x \in k\text{NN}(\varphi \cup \{x\}, z)$ . So again, the addition of  $x$  can only remove such an edge and hence again (2.4) holds for  $y$ .  $\square$

Note that the approach presented above also applies to further graphs studied in computational geometry. The most immediate adaptation concerns the *bidirectional  $k$ -nearest neighbor graph*, see [BB13], where in the definition of the score function, we replace  $\mathfrak{R}_k(\varphi) \vee \mathfrak{R}_k(\varphi - x)$  by  $\mathfrak{R}_k(\varphi) \wedge \mathfrak{R}_k(\varphi - x)$ . Not only can we take the same radius of stabilization, but also Lemma 2.1 remains valid. As a third example, we showcase the *relative neighborhood graph*. Here, for  $\alpha \geq 0$  and  $\varphi \in \mathbf{N}_o$  the score function is given by

$$\xi_{\text{RN}}(\varphi) = \frac{1}{2} \sum_{x \in \varphi: \varphi \cap B_{|x|}(o) \cap B_{|x|}(x) = \emptyset} |x|^\alpha.$$

The *relative neighborhood graph* is a sub-graph of the Delaunay tessellation, and in fact we can reuse the radius of stabilization from Section 2.2. Finally, proving the analog of Lemma 2.1 reduces to the observation that the degree of every node in the relative neighborhood graph is bounded by a constant  $K = K(d)$ , see [JT92, Section IV]. What remains to be verified is that  $\xi_{\text{RN}}$  is weakly decreasing.

**Lemma 2.2.**  $\xi_{\text{RN}}$  is weakly decreasing.

*Proof.* We claim that for any nonequidistant configuration  $\varphi \cup \{x\}$  with  $\varphi \in \mathbf{N}$ , for all but at most  $K$  points  $y \in \varphi$ ,

$$\xi_{\text{RN}}(\varphi \cup \{x\} - y) \leq \xi_{\text{RN}}(\varphi - y) \quad (2.5)$$

holds. Indeed, for  $y \in \varphi$ , let us define the set of relative neighbors of  $y$  in  $\varphi$  as follows

$$\text{RN}(\varphi, y) := \{z \in \varphi \setminus \{y\} : \varphi \cap B_{|z-y|}(y) \cap B_{|z-y|}(z) = \emptyset\},$$

and note that  $z \in \text{RN}(\varphi, y)$  if and only if  $y \in \text{RN}(\varphi, z)$ . In particular,  $\#\text{RN}(\varphi, y) \leq K$  for any  $y \in \varphi$ . So, if  $y \in \text{RN}(\varphi \cup \{x\}, x)$ , then possibly  $\xi_{\text{RN}}(\varphi \cup \{x\} - y) > \xi_{\text{RN}}(\varphi - y)$ . But if  $y \notin \text{RN}(\varphi \cup \{x\}, x)$ , then

$$\begin{aligned} & \xi_{\text{RN}}(\varphi \cup \{x\} - y) - \xi_{\text{RN}}(\varphi - y) \\ &= \frac{1}{2} \sum_{z \in \varphi - y} |z - y|^\alpha \left( \mathbb{1}\{z \in \text{RN}(\varphi \cup \{x\}, y)\} - \mathbb{1}\{z \in \text{RN}(\varphi, y)\} \right) \leq 0, \end{aligned}$$

as asserted.  $\square$

### 3 Proofs

In this section we provide the proofs of the main theorems.

#### 3.1 Proof of Theorem 1.1

The proof of the upper bound relies on the level-3 large deviation principle for the Poisson point process from [GZ93, Theorem 3.1].

*Proof of Theorem 1.1.* Replacing  $\xi^r$  by  $\xi^r \wedge r$  if necessary, we may assume that  $\xi^r$  is bounded above by  $r$ . Then,  $\xi^r$  is a bounded local observable, so that by the contraction principle [DZ98, Theorem 4.2.10] and [GZ93, Theorem 3.1],

$$\limsup_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n \leq a) \leq \limsup_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n^{\xi^r} \leq a) \leq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi^r] \leq a} h(\mathbb{Q}).$$

Hence, it suffices to show that

$$- \lim_{r \uparrow \infty} \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi^r] \leq a} h(\mathbb{Q}) \leq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi] \leq a} h(\mathbb{Q}).$$

Let  $\{\mathbb{Q}_k\}_{k \geq 1}$  be a family of stationary point processes such that  $\mathbb{Q}_k^o[\xi^k] \leq a$  and

$$\lim_{k \uparrow \infty} h(\mathbb{Q}_k) = \lim_{r \uparrow \infty} \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi^r] \leq a} h(\mathbb{Q}).$$

Let  $\mathbb{Q}_*$  be a subsequential limit of  $\{\mathbb{Q}_k\}_{k \geq 1}$ . To simplify the presentation, we may assume  $\mathbb{Q}_*$  to be the limit of  $\{\mathbb{Q}_k\}_{k \geq 1}$ . Then, by monotone convergence,

$$\mathbb{Q}_*^o[\xi] \leq \lim_{r \uparrow \infty} \mathbb{Q}_*^o[\xi^r] = \lim_{r \uparrow \infty} \lim_{k \uparrow \infty} \mathbb{Q}_k^o[\xi^r] \leq \limsup_{k \uparrow \infty} \mathbb{Q}_k^o[\xi^k] \leq a.$$

Since the specific relative entropy  $h$  is lower semicontinuous, we arrive at

$$\liminf_{k \uparrow \infty} h(\mathbb{Q}_k) \geq h(\mathbb{Q}_*) \geq \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi] \leq a} h(\mathbb{Q}),$$

as asserted. □

#### 3.2 Proof of Theorem 1.2

To prove Theorem 1.2, we consider the truncation  $\xi^M = \xi \wedge M$  of the original increasing and  $r$ -dependent score function  $\xi$  at a large threshold  $M > 1$  and write  $H_n^M = H_n^{\xi^M}$ . In comparison to the arguments in Section 3.1, the proof of the lower bound is more involved, since we can no longer replace  $\mathbb{P}(H_n \leq a)$  by  $\mathbb{P}(H_n^M \leq a)$ . Instead, we rely on a sprinkling approach. For this method to work, we need that the total number of points in pathological areas is small with high probability. More precisely, we say that a point  $X_i \in X$  is  $b$ -dense if  $X(Q_r(X_i)) > b$  and write

$$N_{b,n} = N_{b,n}(X) = \#\{X_i \in X \cap Q_n : X_i \text{ is } b\text{-dense}\}$$

for the total number of  $b$ -dense points in  $Q_n$ . Then,  $b$ -dense points are indeed rare.

**Lemma 3.1** (Rareness of  $b$ -dense points). *Let  $\delta > 0$ . Then,*

$$\limsup_{b \uparrow \infty} \limsup_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(N_{b,n} > \delta n^d) = -\infty.$$

In the second step, we remove all  $b$ -dense points through the coupling. That is, we let  $X^{-,\varepsilon}$  be an independent thinning of  $X$  with survival probability  $1 - \varepsilon$ . Furthermore, we let  $X^{+,\varepsilon}$  be an independent Poisson point process with intensity  $\varepsilon > 0$ . Then, the coupled process

$$X^\varepsilon = X^{-,\varepsilon} \cup X^{+,\varepsilon}$$

is again a Poisson point process with intensity 1. Now, let

$$E_{b,n} = \{X^{+,\varepsilon} \cap Q_n = \emptyset\} \cap \{X^{-,\varepsilon} \cap Q_n \text{ has no } b\text{-dense points}\}$$

be the event that  $X^{+,\varepsilon}$  has no points in  $Q_n$  and that  $X^{-,\varepsilon}$  does not contain any  $b$ -dense points in  $Q_n$ .

**Lemma 3.2** (Removal of  $b$ -dense points). *Let  $b, n, \varepsilon > 0$ . Then,  $\mathbb{P}$ -almost surely,*

$$\mathbb{P}(E_{b,n} | X) \geq \exp(-\varepsilon n^d + N_{b,n} \log(\varepsilon)).$$

Before showing Lemmas 3.1 and 3.2, we illustrate how they enter the proof of (1.3).

*Proof of Theorem 1.2.* Let  $M > 0$ . Then, by [GZ93, Theorem 3.1],

$$\liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n^M < a) \geq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi^M] < a} h(\mathbb{Q}) \geq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi] < a} h(\mathbb{Q}).$$

Hence, it remains to show that

$$\liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n < a) \geq \liminf_{M \uparrow \infty} \liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n^M < a). \quad (3.1)$$

Let  $b, \delta, \varepsilon > 0$  be arbitrary. Now, since  $\xi$  is increasing,

$$\mathbb{P}(H_n < a) = \mathbb{P}(H_n(X^\varepsilon) < a) \geq \mathbb{P}(\{H_n^{M(b)} < a\} \cap E_{b,n}) = \mathbb{E}[\mathbb{1}\{H_n^{M(b)} < a\} \mathbb{P}[E_{b,n} | X]].$$

Thus, by Lemma 3.2,

$$\begin{aligned} \mathbb{P}(H_n < a) &\geq \exp(-\varepsilon n^d) \mathbb{E}[\mathbb{1}\{H_n^{M(b)} < a\} \varepsilon^{N_{b,n}}] \\ &\geq \exp((\delta \log(\varepsilon) - \varepsilon)n^d) \mathbb{P}(H_n^{M(b)} < a) - \mathbb{P}(N_{b,n} > \delta n^d). \end{aligned}$$

Since  $X$  and  $X^\varepsilon$  share the same distribution, Lemma 3.1 allows us to choose  $b = b(\delta) > 0$  sufficiently large such that

$$\liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n < a) \geq \delta \log(\varepsilon) - \varepsilon + \liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n^{M(b)} < a).$$

Hence, sending  $\varepsilon \downarrow 0$ ,  $\delta \downarrow 0$ , and  $b \uparrow \infty$  concludes the proof of (3.1).  $\square$

*Proof of Lemma 3.1.* Consider a subdivision of  $Q_n$ , for sufficiently large  $n \geq 1$ , into sub-boxes  $Q_a(z_i) = z_i + Q_a$  of side length  $a > r$  where  $z_i \in a\mathbb{Z}^d$ . Let  $N_i = X(Q_a(z_i))$  be the number of points in the  $i$ th sub-box and  $N'_i = X(Q_{3a}(z_i))$  be the number of points the  $i$ th sub-box plus its adjacent sub-boxes. Then,  $N_{b,n} \leq N''_{b,n}$ , where

$$N''_{b,n} = \sum_{i \in a\mathbb{Z}^d \cap Q_n} N_i \mathbb{1}\{N'_i > b\},$$

so that by the exponential Markov inequality, for all  $t > 0$ ,

$$\log \mathbb{P}(N_{b,n} > \delta n^d) \leq \log \mathbb{P}(N''_{b,n} > \delta n^d) \leq -\delta t n^d + \log \mathbb{E}[\exp(t N''_{b,n})].$$

Since the random variables  $N_i \mathbb{1}\{N'_i > b\}$  and  $N_j \mathbb{1}\{N'_j > b\}$  are independent whenever  $\|z_i - z_j\|_\infty \geq 3$ , we have  $3^d$  regular sub-grids of  $a\mathbb{Z}^d$  containing independent random variables  $N_i \mathbb{1}\{N'_i > b\}$ . Thus, using Hölder's inequality, independence and the dominated convergence theorem, we arrive at

$$\limsup_{b \uparrow \infty} \limsup_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{E}[\exp(t N''_{b,n})] \leq \frac{1}{(3a)^d} \limsup_{b \uparrow \infty} \log \mathbb{E}[\exp(3^d t N_o \mathbb{1}\{N'_o > b\})] = \frac{1}{(3a)^d}.$$

Since  $t > 0$  was arbitrary, we conclude the proof.  $\square$

*Proof of Lemma 3.2.* First, since  $X^{+, \varepsilon}$  and  $X^{-, \varepsilon}$  are independent, it suffices to compute

$$\mathbb{P}(X^{+, \varepsilon} \cap Q_n = \emptyset \mid X) \quad \text{and} \quad \mathbb{P}(X^{-, \varepsilon} \cap Q_n \text{ has no } b\text{-dense points} \mid X)$$

separately. The void probabilities for a Poisson point process give that

$$\mathbb{P}(X^{+, \varepsilon} \cap Q_n = \emptyset \mid X) = \exp(-\varepsilon n^d).$$

Next, since  $X^{-, \varepsilon}$  is an independent thinning of  $X$  with probability  $\varepsilon$ , we arrive at

$$\mathbb{P}(X^{-, \varepsilon} \cap Q_n \text{ has no } b\text{-dense points} \mid X) \geq \varepsilon^{N_{b,n}},$$

as asserted.  $\square$

### 3.3 Proof of Theorem 1.3

In order to prove the lower bound for stabilizing score functions, we use sprinkling to regularize sub-regions that are not sufficiently stabilized. Let us define the approximation

$$\xi^{\delta, M}(\varphi) = \xi(\varphi \cap Q_M) \wedge \delta M^d$$

and write  $H_n^{\delta, M} = H_n^{\xi^{\delta, M}}$ .

Similarly as before, we consider a coupling construction. Now, we let  $X^{-, M}$  denote an independent thinning of  $X$  with survival probability  $1 - M^{-d}$  and  $X^{+, M}$  an independent Poisson point process with intensity  $M^{-d}$ . Then,

$$X^M = X^{-, M} \cup X^{+, M}$$

defines a unit-intensity Poisson point process.

In this coupling, we consider events in which the sprinkling  $X^{+, M}$  adds points wherever necessary to reduce the stabilization radius. More precisely, let

$$E_n^M = \{X^{-, M} \cap Q_n = X \cap Q_n\} \cap \{X^{+, M}(Q_n) \leq K_0(n/M)^d\} \cap E_n^{M, +}.$$

As we shall prove below, the events  $E_n^M$  occur with a high probability.

**Lemma 3.3** (Sprinkling regularizes with high probability). *Let  $\delta > 0$  and  $n \geq M \geq 1$  sufficiently large. Then, under the assumptions of Theorem 1.3,  $\mathbb{P}$ -almost surely,*

$$\mathbb{P}(E_n^M | X) \geq \exp(X(Q_n) \log(1 - M^{-d}) - \delta n^d).$$

*Proof.* Indeed, for given  $X$ , the event  $\{X^{-,M} \cap Q_n = X \cap Q_n\}$  has probability  $(1 - M^{-d})^{X(Q_n)}$  and is independent of the event  $\{X^{+,M}(Q_n) \leq K_0(n/M)^d\} \cap E_n^{M,+}$ , which has probability at least  $\exp(-\delta n^d)$ .  $\square$

Now, we conclude the proof of Theorem 1.3.

*Proof of Theorem 1.3.* Let  $\delta > 0$  and  $M = M(\delta) > 1$  sufficiently large. Then, by  $R$ -boundedness,

$$\mathbb{P}(H_n < a) = \mathbb{P}(H_n(X^M) < a) \geq \mathbb{P}(\{H_n^{\delta,M}(X^M) < a\} \cap E_n^M).$$

Moreover, under the event  $E_n^M$ ,

$$\begin{aligned} H_n^{\delta,M}(X^M) &= \frac{1}{n^d} \sum_{X_i \in X^{+,M} \cap Q_n} \xi^{\delta,M}(X^M - X_i) + \frac{1}{n^d} \sum_{X_i \in X \cap Q_n} \xi^{\delta,M}(X^M - X_i) \\ &\leq K_0 \delta + H_n^{\delta,M}(X) + \frac{1}{n^d} \sum_{X_i \in X \cap Q_n} (\xi^{\delta,M}(X^M - X_i) - \xi^{\delta,M}(X - X_i)). \end{aligned}$$

Let us write  $X^{M,0} = X$  and  $X^{M,j+1} = X^{M,j} \cup \{X_j^{+,M}\}$  where  $\{X_j^{+,M}\}_{1 \leq j \leq N(M)}$  is an arbitrary ordering of  $X^{+,M}$ . Then, since  $\xi$  is weakly decreasing,

$$\begin{aligned} &\sum_{X_i \in X \cap Q_n} (\xi^{\delta,M}(X^M - X_i) - \xi^{\delta,M}(X - X_i)) \\ &= \sum_{X_i \in X \cap Q_n} \sum_{j \leq N(M)} (\xi^{\delta,M}(X^{M,j} - X_i) - \xi^{\delta,M}(X^{M,j-1} - X_i)) \\ &\leq \delta M^d \sum_{j \leq N(M)} \sum_{X_i \in X \cap Q_n} \mathbb{1}\{\xi^{\delta,M}(X^{M,j} - X_i) > \xi^{\delta,M}(X^{M,j-1} - X_i)\} \\ &\leq k \delta M^d N(M). \end{aligned}$$

Further note that  $N(M) \leq K_0(n/M)^d$ , and thus we arrive at

$$\mathbb{P}(H_n(X^M) < a) \geq \mathbb{P}(\{H_n^{\delta,M}(X^M) < a\} \cap E_n^M) \geq \mathbb{P}(\{H_n^{\delta,M}(X) < a - 2kK_0\delta\} \cap E_n^M).$$

Now, by conditioning on  $X$  and applying Lemma 3.3 for sufficiently large  $n \geq M \geq 1$ ,

$$\begin{aligned} \mathbb{P}(H_n(X^M) < a) &\geq \mathbb{E}[\mathbb{1}\{H_n^{\delta,M}(X) < a - 2kK_0\delta\} \mathbb{P}(E_n^M | X)] \\ &\geq \mathbb{E}[\mathbb{1}\{H_n^{\delta,M}(X) < a - 2kK_0\delta\} \exp(X(Q_n) \log(1 - M^{-d}))] \exp(-\delta n^d). \end{aligned}$$

Moreover, for any  $c > 0$ ,

$$\begin{aligned} &\mathbb{E}[\mathbb{1}\{H_n^{\delta,M}(X) < a - 2kK_0\delta\} \exp(X(Q_n) \log(1 - M^{-d}))] \\ &\geq \exp(cn^d \log(1 - M^{-d})) \mathbb{P}(\{H_n^{\delta,M}(X) < a - 2kK_0\delta\} \cap \{X(Q_n) < cn^d\}), \end{aligned}$$

where for the first factor,

$$\liminf_{M \uparrow \infty} \frac{1}{n^d} \log \left( \exp \left( cn^d \log(1 - M^{-d}) \right) \right) = \liminf_{M \uparrow \infty} c \log(1 - M^{-d}) = 0.$$

Now, for the second factor,

$$\mathbb{P}(\{H_n^{\delta,M}(X) < a - 2kK_0\delta\} \cap \{X(Q_n) < cn^d\}) \geq \mathbb{P}(H_n^{\delta,M}(X) < a - 2kK_0\delta) - \mathbb{P}(X(Q_n) \geq cn^d),$$

where for large  $c$  the second summand plays no role in the large deviations. Applying [GZ93, Theorem 3.1] on the local bounded observable  $\xi^{\delta,M}$  yields that

$$\liminf_{n \uparrow \infty} \frac{1}{n^d} \log \mathbb{P}(H_n^{\delta,M}(X) < a - 2kK_0\delta) \geq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi^{\delta,M}] < a - 2kK_0\delta} h(\mathbb{Q}).$$

Finally, if  $\mathbb{Q}^o[\xi] < a$ , then  $\limsup_{M \uparrow \infty} \mathbb{Q}^o[\xi^{\delta,M}] < a - 2kK_0\delta$  for a sufficiently small  $\delta > 0$ , so that

$$\liminf_{M \uparrow \infty} \left( - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi^{\delta,M}] < a - 2kK_0\delta} h(\mathbb{Q}) \right) \geq - \inf_{\mathbb{Q}: \mathbb{Q}^o[\xi] < a} h(\mathbb{Q}),$$

as asserted. □

## References

- [ACC<sup>+</sup>83] M. Aizenman, J. T. Chayes, L. Chayes, J. Fröhlich, and L. Russo. On a sharp transition from area law to perimeter law in a system of random surfaces. *Comm. Math. Phys.*, 92(1):19–69, 1983.
- [BB13] P. Balister and B. Bollobás. Percolation in the  $k$ -nearest neighbor graph. *In Recent Results in Designs and Graphs: a Tribute to Lucia Gionfriddo, Quaderni di Matematica*, 28:83–100, 2013.
- [BCY18] J.-D. Boissonnat, F. Chazal, and M. Yvinec. *Geometric and Topological Inference*. Cambridge University Press, Cambridge, 2018.
- [CH14] S. Chatterjee and M. Harel. Localization in random geometric graphs with too many edges. *arXiv:1401.7577*, 2014.
- [Cha12] S. Chatterjee. The missing log in large deviations for triangle counts. *Random Structures Algorithms*, 40(4):437–451, 2012.
- [DZ98] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, New York, second edition, 1998.
- [ERS15] P. Eichelsbacher, M. Raič, and T. Schreiber. Moderate deviations for stabilizing functionals in geometric probability. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(1):89–128, 2015.
- [GZ93] H.-O. Georgii and H. Zessin. Large deviations and the maximum entropy principle for marked point random fields. *Probab. Theory Related Fields*, 96(2):177–204, 1993.

- [Jan90] S. Janson. Poisson approximation for large deviations. *Random Structures Algorithms*, 1(2):221–229, 1990.
- [JT92] J. W. Jaromczyk and G. T. Toussaint. Relative neighborhood graphs and their relatives. *Proceedings of the IEEE*, 80(9):1502–1517, 1992.
- [JW16] S. Janson and L. Warnke. The lower tail: Poisson approximation revisited. *Random Structures Algorithms*, 48(2):219–246, 2016.
- [Pen07] M. D. Penrose. Gaussian limits for random geometric measures. *Electron. J. Probab.*, 12:989–1035, 2007.
- [SW08] R. Schneider and W. Weil. *Stochastic and Integral Geometry*. Springer, Berlin, 2008.
- [SY01] T. Seppäläinen and J. E. Yukich. Large deviation principles for Euclidean functionals and other nearly additive processes. *Probab. Theory Related Fields*, 120(3):309–345, 2001.
- [SY05] T. Schreiber and J. E. Yukich. Large deviations for functionals of spatial point processes with applications to random packing and spatial graphs. *Stochastic Process. Appl.*, 115(8):1332–1356, 2005.